Convergent, Bounding Approximation Procedures with Applications to the Ferromagnetic Ising Model*

GEORGE A. BAKER, JR.
Applied Mathematics Department, Brookhaven National Laboratory, Upton, New York

(Received 20 April 1967)

The Padé-approximant method is generalized in such a way that converging upper and lower bonds can be established from the early power-series coefficients for a wider class of functions than was previously possible. These procedures are proved applicable to many thermodynamic properties of the ferromagnetic Ising model and used thereon. Certain pitfalls of nonbounding calculational methods, when applied to this problem, are noted.

1. INTRODUCTION AND SUMMARY

In attempting to treat problems for which no simple, closed-form solution is readily available, it is desirable to have systematic approximation procedures which converge rapidly, and, in practical cases, with ascertainable accuracy. One way to develop such a scheme is to base it on a perturbation- or Taylor-series expansion of the quantity sought. In many cases, for reasons not particularly related to the physics of the problem, the series fail to converge for the value of interest. In order to make further headway with this type of approach, methods of series summation or approximate analytic continuation can be developed. For example, the Padé-approximant method[1] has proved very successful in this regard. For a certain class of functions (series of Stieltjes) it provides converging rigorous upper and lower bounds. The principal known physical examples of series of Stieltjes are[1] the forward scattering amplitude for potential scattering and the Ladder approximation to the many-body energy with a purely repulsive potential. For a much larger class of functions, the Padé-approximant method provides rapidly converging estimates, but no indication of the error involved, save for the relative consistency of various different Padé approximants.

It is the purpose of this paper to widen the class of functions for which convergent upper and lower bounds can be given, by developing an extension of the Padé approximants. We then show that this widened class includes various thermodynamic properties of the ferromagnetic Ising model. The value of these various thermodynamic properties can, in principle, be computed for any temperature and (nonzero) magnetic field to arbitrary and precisely definable accuracy, by use of these procedures.

In the second section, we introduce approximants of the form

$$B_{n,j}(x)=\sum_{m=1}^{n}\alpha_n b(x,\sigma_m)+\sum_{k=0}^{j}\frac{\beta_k}{k!}\left[\left(\frac{\partial}{\partial s}\right)^k b(x,s)\bigg|_{s=0}\right], \tag{1.1}$$

and consider their use in approximating functions of the type

$$g(z)=\int_0^\infty b(z,s)\,d\phi(s), \tag{1.2}$$

where $d\phi\geq 0$. We prove several general theorems about this type of approximant. We give a sufficient condition that (1.1) converge, and that it converge to a unique function of z. Among the kernels $b(x,s)$ for which we can prove convergence, we give a simple, necessary, and sufficient condition that (1.1) form upper and lower bounds to $g(z)$, namely,

$$(-\partial/\partial s)^j b(x,s)\geq 0 \tag{1.3}$$

for all real, positive x and s, and $j=0, 1, 2, \cdots$. We establish that the most general function of the form (1.3) which is, in addition, regular in the right half-plane and vanishes sufficiently fast as s tends to infinity, is of the form

$$b(x,s)=\int_0^\infty e^{-st}d\phi(x,t), \tag{1.4}$$

where $d\phi\geq 0$. In the third section, we illustrate these results with a few sample kernels, $b(x,s)$.

In the fourth section, we prove that the free energy of the ferromagnetic Ising model, when considered as a function of

$$\mu=\exp(-2mH/kT) \tag{1.5}$$

can, through these procedures, be given converging upper and lower bounds for every nonzero, magnetic field. The same is true for every derivative with respect to magnetic field, i.e., spontaneous magnetization, magnetic susceptibility, etc.

In the fifth section, we explain how a rigorous, though inconvenient, bound on the energy can be given. We then make a conjecture on the temperature dependence of the roots of the grand partition function, test it experimentally against all available series data, and use it to construct much more convenient bounds for the energy.

In the final section we investigate calculation of the various thermodynamic properties of the Ising model without the use of bounding procedures. We find that

---

* Work performed under the auspices of the U. S. Atomic Energy Commission.

[1] See, for example, G. A. Baker, Jr., Advan. Theoret. Phys. 1, 1 (1965), and references therein.

standard procedures such as Padé approximation and Taylor-series summation work well, except that near the critical point there is deceptively slow convergence while still far from the correct answer. In such a situation, converging bounds are very desirable.

2. GENERAL THEOREMS

Such success as has been had to date in proving convergence theorems concerning Padé approximants[1] has arisen primarily from a representation of the function to be approximated in the form

$$f(z)=\int_0^\infty \frac{d\phi(u)}{1+zu}, \tag{2.1}$$

where $d\phi\geq 0$. The Padé approximation then consists of approximating $d\phi$ by a discrete sum of delta functions whose strengths and locations are determined so that the leading coefficients of the power series expansion of the Padé approximants agree with those of the function being approximated. For functions of the form (2.1) (called series of Stieltjes) all locations and coefficients of the delta functions are non-negative real, and various Padé approximants can be proved to form convergent upper and lower bounds to $f(z)$ for z real and positive. In addition, the approximants converge for any z which is not on the negative real axis.

We seek to relax form (2.1) in the hope of expanding the class of functions for which we can establish convergence theorems and bounds on the errors involved in the use of finite approximations. To this end let us consider

$$g(z)=\int_0^\infty b(z,s)\,d\phi(s), \tag{2.2}$$

where $d\phi\geq 0$ and the properties of $b(z,s)$ are not yet specified. Our approximants would again be defined by approximating $d\phi$ as a sum of delta functions such that the leading power-series coefficients agree with those of $g(z)$. Some other authors[2] have considered special cases of (2.2), but have not, to our knowledge, proved any theorems about them.

We shall now introduce a triangular table of approximants.

$$B_{n,j}(x)=\sum_{m=1}^{n}\alpha_m b(x,\sigma_m)+\sum_{k=0}^{j}\frac{\beta_k}{k!}\left[\left(\frac{\partial}{\partial s}\right)^k b(x,s)\Big|_{s=0}\right], \tag{2.3}$$

where $n=1,2,\cdots,$ and $j=-1,0,1,\cdots.$ These approximants are, for $b(x,s)=1/(1+xs)$, the $[n,n+j]$ Padé approximants. When $j=-1$, the second sum in (2.3) is to be omitted entirely. The defining equations for α, β, σ are obtained as follows: Let us expand

$$b(x,s)=\sum_{m=0}^{\infty} b_n(x)(-s)^m \tag{2.4}$$

and denote

$$c_n=\int_0^\infty s^n d\phi(s). \tag{2.5}$$

Then, equating (2.2) to (2.3), we have

$$\sum_{m=0}^{\infty} b_m(x)\sum_{l=1}^{n}\alpha_l(-\sigma_l)^m+\sum_{m=0}^{j} b_m(x)\beta_m(-1)^m = \sum_{m=1}^{\infty} b_m(x)(-1)^n c_m. \tag{2.6}$$

Let us assume that $b(x,s)$ satisfies a "solvability" condition so that (2.6) can be made equivalent to a set of equations obtained by equating the coefficients of $b_m(x)$. For example, if $b_m(x)\propto x^m$ as x goes to zero, then we could obtain this equivalence. We take therefore the defining equations to be

$$\sum_{l=1}^{n}\alpha_l(\sigma_l)^k+\beta_k=c_k, \qquad 0\leq k\leq j$$

$$\sum_{l=1}^{n}\alpha_l(\sigma_l)^k=c_k, \qquad j<k\leq 2n+j, \tag{2.7}$$

where there are $2n+j+1$ equations in the same number of parameters and the c_k are determined by the solution of

$$g(z)=\sum_{m=0}^{\infty} b_m(z)c_m(-1)^m. \tag{2.8}$$

If we identify the $(-\sigma_l)^{-1}$ as the location of poles and (α_l/σ_l) as the respective residues, then (2.7) are exactly the equations for the $[n,n+j]$ Padé approximants to the function

$$C(z)=\sum_{k=0}^{\infty} c_k(-z)^k. \tag{2.9}$$

Thus, for suitably restricted $b(x,s)$, we may reduce the study of this more general approximation procedure to the study of the Padé approximants to a transformed function. The series $C(z)$ will be a series of Stieltjes for $g(z)$ of the form (2.2).

We are now in a position to give some sufficient conditions on $b(x,s)$ so that the approximants (2.3) will converge to $g(z)$.

Theorem 1. Suppose $b(x,s)$ is regular in a uniform neighborhood of the positive, real s axis and $(\ln s)^{1+\eta}\times b(x,s)$ is bounded as $s\to+\infty$, for some $\eta>0$; then the approximants $B_{n,j}(x)$ converge as n goes to infinity for functions $g(z)$ of the form (2.2).

[2] D. P. Taylor, J. L. Gammel, and C. Rousseau. Bull. Am. Phys. Soc. **12**, 83 (1967); R. D. Teasdale, Institute of Radio Engineers Convention Record, 1953, Part 5, p. 89 (unpublished). I am indebted to Dr. A. V. M. Ferris-Prabhu for this latter reference.

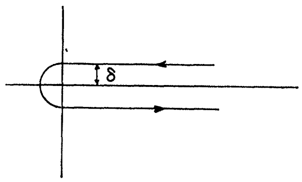

FIG. 1. Integration contour in the s plane for Eq. (2.10).

Proof. Consider

$$(2\pi i)^{-1}\int_C b(x,s)[n,n+j]\left(\frac{-1}{s}\right)\frac{ds}{s}, \qquad (2.10)$$

where the contour C is as shown in Fig. 1. We wish to prove that (2.10) converges as n tends to infinity. By theorem 7 of Ref. 1, $[n,n+j](-1/s)$ converges at every point of C as $n\to\infty$. Furthermore, $s^{-1}[n,n+j]\times(-1/s)$ is bounded at every point of the contour, since $[n,n+j]$ approximates a series of Stieltjes. This result follows easily as the α_n, being positive, are bounded by their sum which is c_0. Let us now break the integral (2.10) into two parts. In one part we consider that portion of C such that $|s|\leq S$ and in the other part $|s|>S$. The second part may be written as

$$1/2\pi i\int_{C>S} b(x,s)\left[\sum_{m=1}^{n}\frac{\alpha_m}{s-\sigma_m}+\sum_{k=0}^{j}\beta_k s^{-k-1}\right]ds. \qquad (2.11)$$

Now, except for s in the range $|s-\sigma_m|\leq\frac{1}{2}\sigma_m$,

$$1/|s-\sigma_m|<3/\sigma. \qquad (2.12)$$

In the range $|s-\sigma_m|\leq\frac{1}{2}\sigma_m$,

$$1/|s-\sigma_m|=1/[(\sigma-\sigma_m)^2+\delta^2]^{1/2}, \qquad (2.13)$$

where σ is the real part of s.

Using (2.12) and (2.13), the result that the absolute value of an integral is less than or equal to the integral of the absolute value, the fact that all α's and β's are positive real, and (2.7) for $k=0$, we may write

$$|(2.11)|\leq\frac{2}{2\pi}\int_S^{\infty}d\sigma\,|b(x,s)|\left[\frac{3c_0}{\sigma}+\sum_{k=1}^{j}\beta_k\sigma^{-k-1}\right]d\sigma$$
$$+\frac{2}{2\pi}\sum_{m=1}^{n}{}'\,\alpha_m\int_{(1/2)\sigma_m}^{(3/2)\sigma_m}d\sigma\,|b(x,s)|[(\sigma-\sigma_m)^2+\delta^2]^{-1/2}, \qquad (2.14)$$

where the primed summation includes only those terms with $\sigma_m>\frac{1}{2}S$.

Now, assuming the bound on $b(x,s)$ stated in the theorem

$$|b(x,s)|<K/(\ln\sigma)^{1+\eta}, \qquad (2.15)$$

we may compute (2.14) as

$$|(2.11)|\leq\frac{6c_0K}{2\pi\eta}(\ln S)^{-\eta}+\frac{2K}{2\pi}(\ln S)^{-1-\eta}\sum_{k=1}^{j}C_jS^{-k}$$
$$+\sum_{m=1}^{n}\alpha_m\frac{2K\ln[\frac{1}{2}\sigma_m+(\frac{1}{4}\sigma_m{}^2+\delta^2)^{1/2}]}{2\pi[\ln(\frac{1}{2}\sigma_m)]^{1+\eta}}. \qquad (2.16)$$

As the function in the last summation is monotonically decreasing in σ_m (as $\sigma_m>\frac{1}{2}S\gg\delta$), we may bound this term by $2Kc_0/(\ln S)^{-\eta}$. Hence

$$|(2.11)|\leq\frac{4c_0K}{\pi\eta}(\ln S)^{-\eta}+\frac{K}{\pi}(\ln S)^{-1-\eta}\sum_{k=1}^{j}c_jS^{-k} \qquad (2.17)$$

uniformly in n, for S and j fixed. Thus, given any error, $\epsilon>0$, we may pick an S such that the bound of (2.17) is less than $\frac{1}{2}\epsilon$. We may then, by theorem 7 of Ref. 1, pick an n_0 such that all $[n,n+j]$ Padé approximants with $n>n_0$ differ by less than

$$\epsilon/[2(S+\pi\delta)\max_{C<S}\{|b(x,s)|\}] \qquad (2.18)$$

at every point of $C<S$, from their limiting value. Consequently the integral (2.10) converges in the limit as n goes to infinity. However, from the form exemplified by (2.11) it follows by Cauchy's integral theorem that (2.10) is exactly $B_{n,j}(x)$. Q.E.D.

Corollary 1. If, in addition,

$$\sum_{p=1}^{\infty}(c_p)^{-1/(2p+1)}$$

diverges, $\lim_{n\to\infty}B_{n,j}(x)$ is independent of j.

Proof. This result follows from Theorem 7 of Ref. 1 and is the condition that all diagonal sequences of Padé approximants converge to a common limit.

Corollary 2. If the series c_p has a finite radius of convergence R then we may replace the conditions on $b(x,s)$ in theorem 1 by $b(x,s)$ regular in the neighborhood of $0\leq s\leq R^{-1}$. (Note: The condition of corollary 1 is automatically satisfied here.)

Proof. As all poles of the Padé approximant $[n,n+j](-1/s)$ lie[1] in the range $0\leq s\leq R^{-1}$, we may choose, instead of the contour C in (2.10), C' which encircles that line segment and on which $b(x,s)$ is regular. As the Padé approximants converge (theorem 7, Ref. 1) at every point of C', they converge uniformly and hence the $B_{n,j}(x)$ converge.

One of the advantageous results for Padé approximants to series of Stieltjes is that they form monotonically converging (on the positive real axis) sequences which give upper and lower bounds to the correct answer, and thus allow assessment of the error of the approximants. As the next theorem will show, a remarkably simple, additional property of the function $b(x,s)$ is both necessary and sufficient to ensure the same bounding properties for sequence of $B_{n,j}(x)$ approximants.

Theorem 2. The approximants $B_{n,j}(x)$ to a function of the form (2.2) obey the following inequalities where x is real and non-negative:

$$(-1)^{1+j}\{B_{n+1,j}(x)-B_{n,j}(x)\}\geq 0, \qquad (2.19a)$$

$$(-1)^{1+j}\{B_{n,j}(x)-B_{n-1,j+2}(x)\}\geq 0, \qquad (2.19b)$$

$$B_{n,0}(x)\geq g(x)\geq B_{n,-1}(x), \qquad (2.19c)$$

where $j \geq -1$, if and only if

$$(-\partial/\partial s)^j b(x, s) \geq 0 \qquad (2.20)$$

for all real, non-negative x and s, and $j=0, 1, 2, \cdots$. These inequalities have the consequence that the $B_{n,0}(x)$ and $B_{n,-1}(x)$ sequences form the best upper and lower bounds obtainable from the $B_{n,j}(x)$ approximants with a given number of coefficients and that the use of additional coefficients (higher n) improves the bounds.

Proof. The first step in our proof is to establish a representation for the inequalities in terms of partial derivatives of b. To this end let us define

$$\frac{\Delta^n f(x_i)}{(n!)} = \frac{\det\begin{vmatrix} 1 & x_i & x_i^2 & \cdots & x_i^{n-1} & f(x_i) \\ 1 & x_{i+1} & x_{i+1}^2 & \cdots & x_{i+1}^{n-1} & f(x_{i+1}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{i+n} & x_{i+n}^2 & \cdots & x_{i+n}^{n-1} & f(x_{i+n}) \end{vmatrix}}{\det\begin{vmatrix} 1 & x_i & x_i^2 & \cdots & x_i^{n} \\ 1 & x_{i+1} & x_{i+1}^2 & \cdots & x_{i+1}^{n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{i+n} & x_{i+n}^2 & \cdots & x_{i+n}^{n} \end{vmatrix}} \qquad (2.21)$$

on the distinct points $(x_i, x_{i+1}, \cdots, x_{i+n})$. It follows easily that

$$\begin{aligned} \Delta^n(x^j) &= 0, \qquad j<n \\ &= n!, \qquad j=n \end{aligned} \qquad (2.22)$$

as for $j<n$ two columns of the numerator are equal and for $j=n$ the numerator and denominator of (2.21) cancel. Thus we identify Δ^n as the nth difference operator.[3]
One can easily establish the usual mean-value-theorem result

$$\Delta^n f(x_i) = f^{(n)}(x_i + \theta(x_{i+n} - x_i)), \qquad 0<\theta<1 \qquad (2.23)$$

by considering

$$P(x) = \det\begin{vmatrix} 1 & x & x^2 & \cdots & x^n & f(x) \\ 1 & x_i & x_i^2 & \cdots & x_i^n & f(x_i) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{i+n} & x_{i+n}^2 & \cdots & x_{i+n}^n & f(x_{i+n}) \end{vmatrix}. \qquad (2.24)$$

This function vanishes for $x=(x_i, x_{i+1}, \cdots, x_{i+n})$ and so by Rolle's theorem[4] $P'(x)$ must vanish for

$$x = (x_i^{(1)}, x_{i+1}^{(1)}, \cdots, x_{i+n-1}^{(1)}),$$

where

$$x_i < x_i^{(1)} < x_{i+1} < x_{i+1}^{(1)} < x_{i+2} < \cdots < x_{i+n-1}^{(1)} < x_{i+n}. \qquad (2.25)$$

If we repeat this argument we eventually obtain $P^{(n)}(x_i^{(n)})=0$ for $x_i < x_i^{(n)} < x_{i+n}$. If we express this result by directly differentiating (2.24), we get (2.23).

Let us now consider inequality (2.19a) for $j=0$. It is, by (2.3),

$$\hat{\beta}_0 b_0(x) + \sum_{l=1}^{n-1} \hat{\alpha}_l b(x, \hat{\sigma}_l) - \beta_0 b_0(x) - \sum_{l=1}^{n} \alpha_l b(x, \sigma_l) = \Omega(b), \qquad (2.26)$$

a functional of b. The points σ_l and $\hat{\sigma}_l$ are all distinct by theorem 5 of Ref. 1. They lie in the order $(0, \sigma_n, \hat{\sigma}_{n-1}, \sigma_{n-1}, \cdots, \hat{\sigma}_1, \sigma_1)$. We observe that

$$\Omega(s^j) = 0, \qquad \text{if } j=0, 1, \cdots, 2n-2 \qquad (2.27)$$

by the fundamental equations (2.7) as both $B_{n,0}(x)$ and $B_{n-1,0}(x)$ approximate the same series. As (2.26) is a $2n$-point formula, it is thus a multiple of Δ^{2n-1}. The coefficient can be obtained from Eq. (II.18) of Ref. 1 as

$$\Omega \equiv -\frac{D(1, n-1)\Delta^{(2n-1)}}{D(1, n-2)(2n-1)!}, \qquad (2.28)$$

since by (2.7) $\Omega(s^{2n-1})$ is $(-1)^{2n-1}$ times the $(2n-1)$st series coefficient in the difference, where

$$D(j, k) = \det\begin{vmatrix} c_j & c_{j+1} & \cdots & c_{j+k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ c_{j+k} & c_{j+k+1} & \cdots & c_{j+2k} \end{vmatrix}. \qquad (2.29)$$

The $D(j, k)$ are all positive [Eqs. (II.10)–(II.11), Ref. 1] as $d\phi \geq 0$ in (2.2). Thus combining our results we have shown that there exists a σ such that

$$(2n-1)![B_{n-1,0}(x) - B_{n,0}(x)] = -\frac{D(1, n-1)}{D(1, n-2)} \frac{\partial^{2n-1}}{\partial s^{2n-1}} b(x, s)\Big|_{s=\sigma}, \qquad (2.30)$$

where $0<\sigma<\sigma_1$. Hence the derivative condition implies the inequalities. The other inequalities follow in exactly the same way as the one we have just proven.

The converse may be shown as follows: In the same way as before

$$(2n-2)![B_{n,-1}(x) - B_{n-1,1}(x)] = \frac{D(0, n-1)}{D(0, n-2)} \frac{\partial^{2n-2} b(x, s)}{\partial s^{2n-2}}\Big|_{s=\sigma}, \qquad (2.31)$$

[3] This definition can be shown to be equivalent to the more usual form, which is given, for example, by H. Jefferies and B. S. Jefferies, *Methods of Mathematical Physics* (Cambridge University Press, New York, 1950), Sec. 9.012.

[4] See, for example, P. Franklin, *A Treatise on Advanced Calculus* (John Wiley & Sons, Inc., New York, 1940), Sec. 73.

where now $\sigma_m < \sigma < \sigma_1$, as 0 is not one of the set of basic points. However, we may choose these arbitrarily by selecting $d\phi$ as a sum of delta functions and so, by making (σ_n, σ_1) narrow and sweeping it past any desired point, we may, by Bolzano's theorem, find an example which has σ at any point we desire. Thus, the inequalities imply the derivative conditions for even-order derivatives. Since $b(x, s)$ goes to zero as $s \to \infty$, it follows that the odd derivatives are of fixed negative sign. This result can be seen as follows: Since $\partial^2 b/\partial s^2$ is positive, if $\partial b/\partial s$ ever becomes positive, it cannot decrease in magnitude and therefore $b(x, s)$ diverges at least linearly as s goes to infinity. But $b(x, s)$ in fact goes to zero as s goes to infinity. Thus $\partial b/\partial s$ is non-positive definite. It goes to zero as s goes to infinity in order that b remain non-negative. Similarly, $\partial^2 b/\partial s^2$ must go to zero. We can now repeat this argument for $\partial^2 b/\partial s^2$, $\partial^3 b/\partial s^3$, $\partial^4 b/\partial s^4$ and thus by induction all odd derivatives of b are nonpositive definite. Consequently the inequalities imply the derivative condition as well as vice versa. Q.E.D.

One may inquire what class of functions is characterized by property (2.20). At least for functions which have certain regularity and boundedness conditions, this class of functions is easily characterized. The following theorem does so and shows that (2.20) characterizes a large class of functions.

Theorem 3. In the class of functions which are regular in the right-half s plane and go to zero faster than s^{-k}, $k > 1$ for some k as s goes to infinity therein the following statements are equivalent:

$$(-\partial/\partial s)^j f(s) \geq 0, \qquad 0 \leq s < +\infty \qquad (2.32)$$

and

$$f(s) = \int_0^\infty e^{-st}\, d\phi(t), \qquad (2.33)$$

with $d\phi \geq 0$.

Proof. If (2.33) holds for $0 \leq s < +\infty$ we may differentiate under the integral sign $0 < s < +\infty$ and consequently obtain (2.32), as $d\phi \geq 0$. (The result for $s = 0$ follows by taking the limit from $s > 0$.)

Suppose now that (2.32) holds. The conditions in the theorem are sufficient[5] to ensure that $f(s)$ can be written as a Laplace transform:

$$f(s) = \int_0^\infty e^{-st} F(t)\, dt, \qquad (2.34)$$

where F is continuous and of order $0(1)$ for all $t \geq 0$. Consider

$$(-\partial/\partial s)^j f(s) = \int_0^\infty t^j e^{-st} F(t)\, dt \qquad (2.35)$$

in the limit where s and j go to infinity together in such a way that $j/s = \tau$, a constant. The function $t^j e^{-st}$ can then have an arbitrarily sharp peak at $t = \tau$ for large enough s and j; thus we get approximately

$$(-\partial/\partial s)^j f(s) \approx A \int \delta(t-\tau) F(t)\, dt \approx A F(\tau) \geq 0 \qquad (2.36)$$

by (2.32). Hence (2.32) also implies (2.33). Q.E.D.

Corollary 3. In the class of functions which are regular in the circle with diameter, $0 \leq s \leq S$, and go to zero faster than $(S-s)^k$, $k > 1$ for some k as s goes to S from within, the following statements are equivalent:

$$(-\partial/\partial s)^j F(s) \geq 0, \qquad 0 \leq s \leq S \qquad (2.37)$$

and

$$F(s) = \int_0^\infty \exp\left[-st \Big/ \left(1 - \frac{s}{S}\right)\right] d\phi(t), \qquad (2.38)$$

with $d\phi \geq 0$.

Proof. If $F(s)$ satisfies (2.37), then

$$f(s) = F\left(\frac{s}{1+s/S}\right) \qquad (2.39)$$

satisfies (2.32). This result follows because

$$(d/ds)^j \left(\frac{s}{1+s/S}\right) = (-S)^{1-j} j! \left(1 + \frac{s}{S}\right)^{-1-j}. \qquad (2.40)$$

Consequently the jth derivative of f is a sum of terms

$$(-d/ds)^j f(s) = \sum \alpha_{r,u} F^{(r)} \left(\frac{s}{(1+s/S)}\right) \left(1 + \frac{s}{S}\right)^u. \qquad (2.41)$$

That every α is positive follows by induction on j as no matter whether $F^{(r)}$ or $(1+s/S)^u$ is differentiated, the sign factor is the same. Repeated differentiation of (2.38) shows that it implies (2.37). The corollary now follows from theorem 3 as the transform (2.39) maps its conditions into those of theorem 3.

These results provide the ground work for applications in which we can construct convergent bounding approximants to a wider class of functions than was previously possible.

3. SAMPLE KERNELS

Before proceeding to the applications, we will mention briefly a few examples which fall in the scope of the results of the previous section. The first is obtained by taking

$$b(z, s) = L_\zeta(zs), \qquad (3.1)$$

where $L_\zeta(x)$ is the LeRoy function[6]

$$L_\zeta(x) = \sum_{n=0}^\infty \frac{\Gamma(1+\zeta n)}{\Gamma(1+n)} (-x)^n, \qquad 0 \leq \zeta < 1. \qquad (3.2)$$

Most of its properties can be readily established from

[5] See, for example, R. V. Churchill, *Modern Operational Mathematics in Engineering* (McGraw-Hill Book Company, Inc., New York, 1944), Sec. 56.

[6] See, G. H. Hardy, *Divergent Series* (Oxford University Press, London, 1956), p. 197.

its integral representation

$$L_\zeta(x)=\int_0^\infty \exp[-(t+xt^\zeta)]dt. \quad (3.3)$$

Repeated differentiation of (3.3) immediately establishes property (2.20). As one may easily show that

$$|L_\zeta(x)|<\Gamma(1/\zeta+1)x^{-1/\zeta} \quad (3.4)$$

from (3.3) for x real and positive, it follows that the conditions for theorems 1 and 2 hold. We see from (3.2) that $L_\zeta(x)$ is an entire function for $0\leq\zeta<1$. The two special cases

$$L_1(x)=(1+x)^{-1}, \qquad L_0(x)=e^{-x} \quad (3.5)$$

are what give this family its interest, for it interpolates by means of entire functions between Padé approximants ($\zeta=1$) and the exponential approximants shown by theorem 3 to be the most general form of kernel of the class $b(s,z)=b(sz)$ for functions regular, etc., in the right-half plane.

One can further show by contour distortion in (3.3) that $|L_\zeta(x)|$ is bounded for the infinite angular wedge defined by $|\arg(x)|\leq\frac{1}{2}\pi(1+\zeta)(1-\epsilon)$, $\epsilon>0$. In the remaining wedge we estimate by the saddle-point method that

$$L_\zeta(x)\sim\exp\{[(\zeta)^{\zeta/(1-\zeta)}-(\zeta)^{1/(1-\zeta)}](-x)^{1/(1-\zeta)}\},$$
$$\tfrac{1}{2}\pi(1+\zeta)<|\arg(x)|\leq\pi. \quad (3.6)$$

One might hope that, since one of the primary difficulties in proving convergence of sequences of Padé approximants has been the occurrence of poles of very small residue in regions of convergence, as the functions $L_\zeta(x)$ are entire, and hence finite for all finite values of their argument, and since in practice the residues are observed to tend to zero very fast, it may be that the whole diagonal sequence of approximants based on $L_\zeta(x)$, $\zeta<1$, converges under very general assumptions rather than just a subsequence, as has been conjectured for Padé approximants.[7]

Another family of kernels is given by

$$b(z,s)=[1+zs/(n+1)]^{-n}, \qquad 0<n<\infty \quad (3.7)$$

which we will use in the next section. The range $n=1$ to ∞ again interpolates between the Padé approximants and the exponential approximants of theorem 3. Repeated differentiation shows directly that (2.20) is satisfied. The conditions of theorem 1 follow as $(\ln s)^{1+n}$ diverges more slowly than any power. Hence theorems 1 and 2 apply to this class of examples as well. Theorem 3 applies, *a fortiori*, as (3.7) is regular and bounded in the whole cut plane ($n>1$). (The case, $n<1$, can also be treated, as is well known.[8]) In addition to the use we make of these in the next section, they provide an alternate method to the standard Padé method[1] of estimating the strength and location of a singularity of known index of divergence, even in the absence of an expansion of the form (2.2).

As one final example, we point out that if $b_1(x,s)$ and $b_2(x,s)$ satisfy (2.20), then so does $b_1(x,s)b_2(x,s)$ by Leibnitz's rule of differentiation.

4. FERROMAGNETIC ISING MODEL

The starting point of our application of the results we have obtained is the result of Yang and Lee.[9] They consider the grand partition function for a lattice gas (equivalent to the partition function for the Ising model) and prove a theorem concerning the location of its zeros as a function of fugacity under the assumption of an impermeable hard core and attractive forces (equivalent to the Ising model with purely ferromagnetic interactions of unspecified range). As they point out, the grand partition function may be written

$$Z_v=\sum_{N=0}^{M}\frac{Q_N}{N!}y^N, \quad (4.1)$$

where "M is the maximum number of atoms that can be crammed into V" on account of the hard cores. As Z_v is a finite polynomial, it can be factored as

$$Z_v=\prod_{i=1}^{M}(1-y/y_i), \quad (4.2)$$

where the $y_1,\cdots,y_M$ are the roots of the algebraic equation

$$Z_v(y)=0. \quad (4.3)$$

None of the roots can be real and positive as the Q_N are all positive. In the mathematically equivalent Ising model,[10] the fugacity is proportional to the variable

$$\mu=\exp(-2mH/kT), \quad (4.4)$$

where m is the magnetic moment, H the magnetic field strength, and T the temperature. Rewriting (4.2) in terms of μ

$$Z_v=\prod_{i=1}^{M}(1-\mu/\mu_i). \quad (4.5)$$

The theorem of Yang and Lee[9] states

$$|\mu_i|=1 \quad (4.6)$$

for all ferromagnetic-type interaction (the spins have lower energy when parallel than antiparallel) Ising models. If we now calculate the spontaneous magneti-

[7] G. A. Baker, Jr., J. L. Gammel, and J. G. Wills, J. Math. Anal. Appl. **2**, 405 (1961); also Ref. 1, and J. S. R. Chisholm, J. Math. Phys. **7**, 39 (1966).

[8] See, for example, G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications* (D. Van Nostrand Company, Inc., New York, 1948) pair 524.

[9] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404, 410 (1952).

[10] For a review of the Ising model see, C. Domb, Phil. Mag. Suppl. **9**, No. 34 (1960); *ibid.* p. 149. M. E. Fisher, J. Math. Phys. **4**, 278 (1963); Rept. Progr. Phys. (to be published).

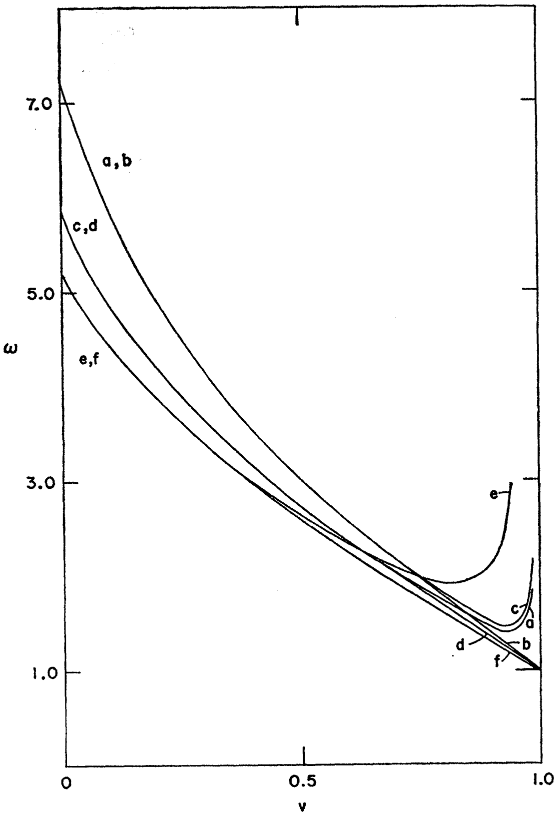

FIG. 2. Upper ($j=0$) and lower ($j=-1$) bounds to $\omega(w)$ at $T=2T_c$ for two-dimensional lattices, plotted against $v\equiv\frac{1}{4}w(1+\frac{1}{4}w)^{-1}$. The large tick on the right-hand side corresponds to the exactly known value $\delta=1$. The curves are (a) honeycomb ($j=0$), (b) honeycomb ($j=-1$), (c) square ($j=0$), (d) square ($j=-1$), (e) triangular ($j=0$), and (f) triangular ($j=-1$).

zation per spin, we have

$$I(\mu)/(Nm)=1-2\mu(d/d\mu)\ln Z$$
$$=\sum_{i=1}^{M}\frac{(1+\mu/\mu_i)}{(1-\mu/\mu_i)}. \quad (4.7)$$

This result means that if $|\mu|\leq 1$ (note the equality sign),

$$\mathrm{Re}[I(\mu)/Nm]\geq 0. \quad (4.8)$$

Consequently, by (II.60) of Ref. 1 we have

$$F(w)=(1+w)^{-1/2}I([(1+w)^{1/2}-1]/[(1+w)^{1/2}+1]), \quad (4.9)$$

which is a series of Stieltjes for all temperatures. This result is perfectly rigorous and without additional assumptions of any kind. Theorems 6 and 7 of Ref. 1 now assure us that the $[N,N](w)$ and $[N,N-1](w)$ Padé approximants form upper and lower bounds, respectively, to $F(w)$ and hence provide them for $I(\mu)$ over the range $0\leq\mu<1$ $(0\leq w<\infty)$. These bounds must converge $0\leq\mu<1$ as $I(\mu)$ is a convergent series with radius 1. The range $1<\mu\leq\infty$ can be computed from the symmetry of the model between points μ and $1/\mu$. For $T>T_c$, the critical temperature (two- or higher-dimensional models), we know that $I(1)=0$, and as the $[N,N-1]$ Padé approximants also equal zero at this point (the $[N,N]$ are infinite here), we have the consequence that we have converging bounds to I everywhere in the (H,T) plane, except on $H=0$, $T<T_c$ (the coexistence curve, for the analogous lattice gas).

Through the use of the results of Sec. 2 we can construct bounding approximations to a number of other thermodynamic properties. According to Yang and Lee,[9] the free energy (related to the logarithm of the partition function) is given by

$$\frac{(F+mH)}{kT}=-\int_0^{2\pi}\ln(1-2\mu\cos\theta+\mu^2)g(\theta)d\theta, \quad (4.10)$$

where $g(\theta)$ is non-negative definite. Letting $\cos\theta=1-2y$ we may rewrite (4.10) as

$$\frac{(F+mH)}{kT}=-\int_0^{1}\ln[(1-\mu)^2+4\mu y]d\phi(y). \quad (4.11)$$

One easily establishes for the allowed range of μ and y that

$$g(\mu,y)=-\ln((1-\mu)^2+4\mu y) \quad (4.12)$$

possesses property (2.20). Hence by corollary 2 and theorem 2 the approximants (2.3) bound the free energy,

$$B_{n,-1}\leq(F+mH)/kT\leq B_{n,0}, \quad (4.13)$$

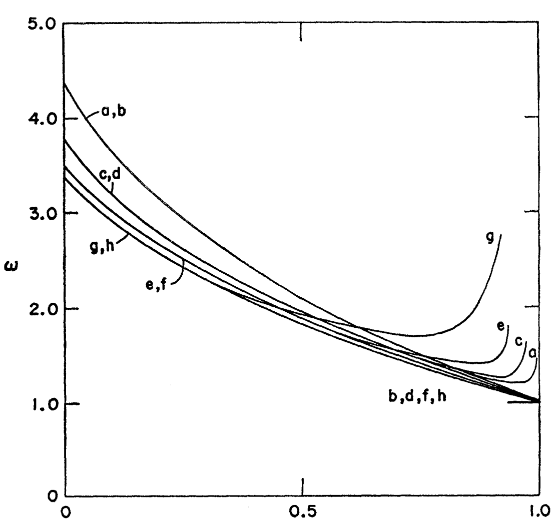

FIG. 3. Upper ($j=0$) and lower ($j=-1$) bounds to $\omega(w)$ at $T=2T_c$ for three-dimensional lattices, plotted against $v\equiv\frac{1}{4}w(1+\frac{1}{4}w)^{-1}$. The large tick on the right-hand side corresponds to the exactly known value $\delta=1$. The curves are (a) diamond ($j=0$), (b) diamond ($j=-1$), (c) simple cubic ($j=0$), (d) simple cubic ($j=-1$), (e) body-centered cubic ($j=0$), (f) body-centered cubic ($j=-1$), (g) face-centered cubic ($j=0$), and (h) face-centered cubic ($j=-1$).

and converge to it [except $B_{n,0}(1)=\infty$] for all temperatures.

The magnetic susceptibility is $2\mu(\partial/\partial\mu)I(\mu)$ and hence is directly related to

$$(d/dw)F(w)=\int\frac{d\phi(\mu)}{(1+w\mu)^2}, \tag{4.14}$$

which is of the type (3.7) with $n=2$, and consequently can be bounded in the same way as the magnetization was, but using a different kernel. All higher derivatives with respect to magnetic field give the sum of successively higher derivatives with respect to w and these terms can be bounded using the same procedures. Since one can bound (above and below) each derivative with respect to w, one can bound a sum (or difference) of such terms.

We can prove a special result about the logarithmic derivative of the magnetization. This quantity is of interest in the estimation of the critical index δ which is defined by the hypothesis that

$$I\propto H^{1/\delta}. \tag{4.15}$$

Gaunt, Fisher, Sykes, and Essam[11] have used it to estimate δ from

$$(1-\mu)(\partial \ln I/\partial\mu)|_{\mu=1}. \tag{4.16}$$

We will now show that we can provide upper and lower bounds to $(\partial \ln I)/(\partial\mu)$. The function $F(w)$, which is directly related to $I(\mu)$ by (4.9), is a series of Stieltjes. Now a series of Stieltjes can be approximated by a converging sequence of $[N,N]$ Padé approximants. Hence

$$\lim_{N\to\infty}\frac{d\ln[N,N]}{dw}=\frac{d\ln F(w)}{dw} \tag{4.17}$$

as the $[N,N]$ are analytic functions in the cut w plane. However,

$$\frac{d\ln[N,N]}{dw}=\sum_{i=1}^{N}[(w+x_i)^{-1}-(w+y_i)^{-1}], \tag{4.18}$$

where the x_i are the poles and the y_i the zeros of the $[N,N]$ Padé approximants. By theorems 5 and 7 of Ref. 1, $y_{i+1}>x_{i+1}>y_i>x_i>R$, where R is the radius of convergence which is unity in this case. We may write the term enclosed by parentheses in (4.18) as

$$\int_{y_i^{-1}}^{x_i^{-1}}\frac{\xi d\xi}{(1+w\xi)^2}=(w+x_i)^{-1}-(w+y_i)^{-1}. \tag{4.19}$$

Hence, if we choose

$$\begin{aligned} d\phi_N(\xi)&=\xi d\xi, && y_i^{-1}<\xi<x_i^{-1}\\ &=0, && \text{otherwise,} \end{aligned} \tag{4.20}$$

one can easily show that at least a subsequence of the ϕ_N converges. They are bounded at every point and hence we may select a countable, dense set on the interval $(0,R^{-1})$. We may now choose a subsequence which converges at the first point; from that, a sub-

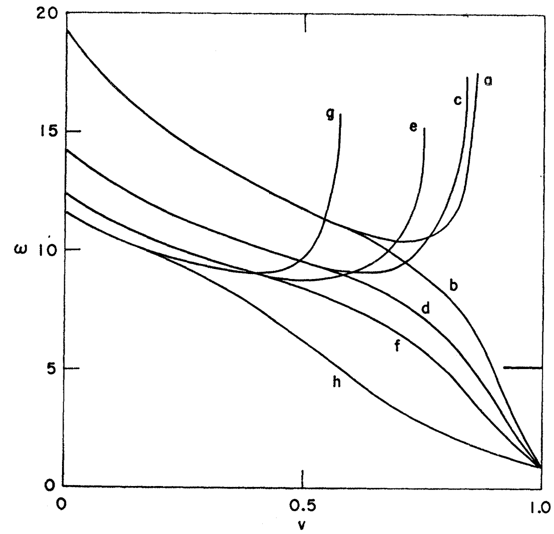

FIG. 5. Upper ($j=0$) and lower ($j=-1$) bounds to $\omega(w)$ at $T=T_c$ for three-dimensional lattices, plotted against $v\equiv\frac{1}{4}w(1-\frac{1}{4}w)^{-1}$. The large tick on the right-hand side corresponds to the accepted value $\delta=5.2$. The curves are (a) diamond ($j=0$), (b) diamond ($j=-1$), (c) simple cubic ($j=0$), (d) simple cubic ($j=-1$), (e) body-centered cubic ($j=0$), (f) body-centered cubic ($j=-1$), (g) face-centered cubic ($j=0$), and (h) face-centered cubic ($j=-1$).

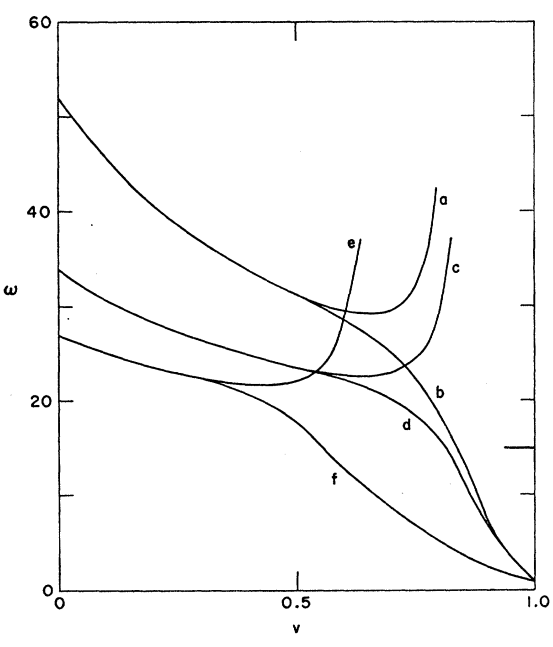

FIG. 4. Upper ($j=0$) and lower ($j=-1$) bounds to $\omega(w)$ at $T=T_c$ for two-dimensional lattices, plotted against $v\equiv\frac{1}{4}w(1+\frac{1}{4}w)^{-1}$. The large tick on the right-hand side corresponds to the accepted value of $\delta=15$. The curves are (a) honeycomb ($j=0$), (b) honeycomb ($j=-1$), (c) square ($j=0$), (d) square ($j=-1$), (e) triangular ($j=0$), and (f) triangular ($j=-1$).

[11] D. S. Gaunt, M. E. Fisher, M. F. Sykes, and J. W. Essam, Phys. Rev. Letters 13, 713 (1964).

sequence which converges at the second point; and so on. However, every ϕ_N is a positive measure, and hence, there exists a limiting ϕ such that

$$\frac{d \ln F(w)}{dw} = \int_0^{R^{-1}} \frac{d\phi(\xi)}{(1+w\xi)^2}, \qquad (4.21)$$

where $d\phi \geq 0$. Consequently we may bound from above and below $d \ln F(w)$ and hence, a function can be constructed which tends to δ which we can also bound $(0 \leq \mu < 1)$.

An additional special result can be given for the logarithmic derivative of a series of Stieltjes. Taking account of the signs, we obtain, by dividing (II.19d) by (II.19c) of Ref. 1,

$$-\frac{d \ln[N, N-1]}{dw} \geq -\frac{d \ln F(w)}{dw} \geq -\frac{d \ln[N, N]}{dw}. \qquad (4.22)$$

To illustrate these results we have computed approximants based on (4.21) and (4.22) for the nearest-neighbor, spin-$\frac{1}{2}$, ferromagnetic Ising model from the extensive data tabulated by Sykes, Essam, and Gaunt.[12] It is worth noting, since these calculations form a sensitive check on the coefficients, that *no* errors were detected. For this case it turns out that (4.22) provides superior bounds to those obtained from (4.21). We have plotted (for $j=0, -1$) in Figs. 2–5

$$\omega(w) = -\tfrac{1}{2}\left[\frac{d \ln[N, N+j]}{dw}(1+w)+\tfrac{1}{2}\right]^{-1}, \qquad (4.23)$$

$$\lim_{w\to\infty} \omega(w) = \delta,$$

versus $v \equiv \frac{1}{4}w(1+\frac{1}{4}w)^{-1}$ so that we expect ω to tend linearly to δ. For $j=-1$, ω always takes on the value 1 for $w=\infty$, which is correct for $T>T_c$ but a lower bound for $T \leq T_c$. For $j=0$, ω always takes on the value -1 for $w=\infty$. Thus the limiting range of the curves (4.23) gives $|\delta| \geq 1$. It will be observed that although these curves form rigorous bounds, their rate of convergence near the critical point is inferior to that found by Gaunt, Fisher, Sykes, and Essam[11] using standard Padé techniques. The point on the right side of the curves is (a) the exactly known value of $\delta=1$ for $T=2T_c$ and (b) the accepted value of δ,[11,13] 15 for two-dimensional lattices and 5.2 for three-dimensional lattices for $T=T_c$. We remark that for $T=\infty$, $\omega(w) \equiv 1.0$, for all lattices.

5. A CONJECTURAL BOUND ON THE ENERGY OF THE FERROMAGNETIC ISING MODEL

In the previous section we showed how all derivatives of the free energy with respect to magnetic field could be bounded from above and below with converging approximants. In this section, we discuss the derivative with respect to temperature. First, let us note that rigorous converging, though somewhat unhandy, bounds can be given as

$$f'(\xi) = \frac{f(a+h)-f(a)}{h} \leq \frac{f_u(a+h)-f_L(a)}{h}, \qquad (5.1)$$

where u means an upper and L a lower bound. The mean-value theorem implies $a<\xi<a+h$.

By convexity, we may use this procedure to bound the derivative of the free energy and thus the energy at any point where we have the bounds of Sec. 4. Bounds for all higher derivatives can also be constructed although we can only specify a range for the point.

A somewhat more convenient bound seems to be valid and may be obtained as follows: Lee and Yang[9] compute the location of all the roots of the grand partition function for the one-dimensional Ising model explicitly. They also give the density of roots for $\mu = \pm 1$ for the two-dimensional Ising model. They observe that the roots (one dimension) move from $\mu=-1$ towards $\mu=+1$ monotonically along the unit circle. Furthermore, for the two-dimensional Ising model the density of roots decreases monotonically at $\mu=-1$ and increases monotonically at $\mu=+1$. The same is also certainly true of the density (spontaneous magnetization) of roots at $\mu=+1$ for the three-dimensional Ising model. It is therefore tempting to conjecture that the roots of the grand partition function move monotonically from $\mu=-1$ to $\mu=+1$ along the unit circle as temperature decreases ($\beta=1/kT$ increases).

In order to use this conjecture, we note the standard statistical mechanical result

$$-(\partial \ln Z)/\partial\beta|_{H\beta} = (E-mH)N = \mathcal{E}N, \qquad (5.2)$$

where E is the total energy per spin and $\mathcal{E}$ is the "configurational" energy per spin. From (4.5) we may rewrite

$$\ln Z = \sum_{j=1}^{M/2} \ln(1-2x_j\mu+\mu^2), \qquad (5.3)$$

where account has been taken that the μ_i appear in complex conjugate pairs and x_j denotes the cosine of the angular position of the root on the unit circle. Therefore

$$\mathcal{E}N = \sum_{j=1}^{M/2} \frac{2\mu(\partial x_j/\partial\beta)}{1-2\mu x_j+\mu^2} \qquad (5.4)$$

$$= \frac{\mu}{1-\mu^2}\left[\frac{1+\mu}{1-\mu}\sum_{j=1}^{M/2} \frac{\partial x_j/\partial\beta}{1+4\mu y_j(1-\mu)^{-2}}\right], \qquad (5.5)$$

where we have substituted $x_j = 1-2y_j$ $(0 \leq y_j \leq 1)$. By theorem 8 of Ref. 1, we see that the term of (5.5) within brackets, if $(\partial x_j)/(\partial\beta) \geq 0$, is exactly of the form of an analytic function which has positive a real part for $|\mu|<1$, and is real for real values of μ. If we apply

[12] M. F. Sykes, J. W. Essam, and D. S. Gaunt, J. Math. Phys. **6**, 283 (1965), especially Appendix III.

[13] G. A. Baker, Jr., and D. S. Gaunt, Phys. Rev. **155**, 545 (1967).

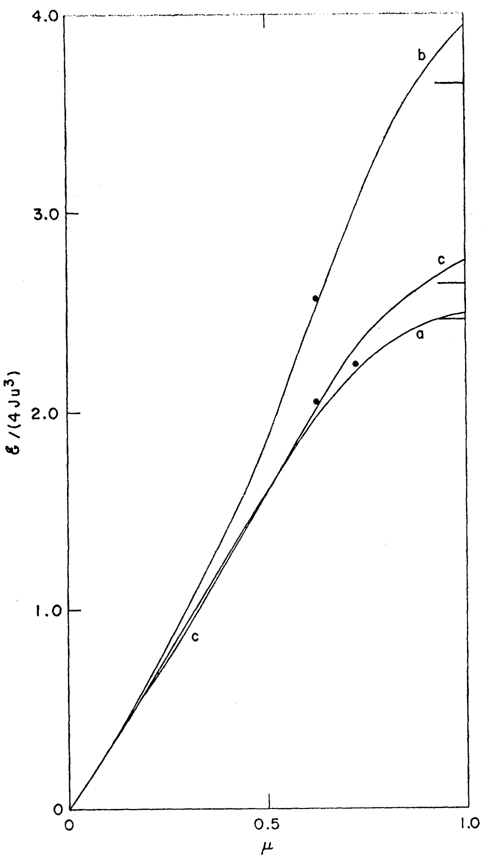

FIG. 6. Lower bounds to $\mathcal{E}/(4Ju^3)$ for the simple-cubic lattice based on the [5, 4] Padé approximant to (5.5). The dots above the curves represent the upper bounds based on the [5, 5] shortly after it becomes distinguishable from the [5, 4]. The ticks below at the right-hand side are the previous lower bound based on the [4, 3]. The curves are: (a) $T=2T_c$ (the slope at $\mu=1$ is small, but not zero, as it is zero for E but not $\mathcal{E}$); (b) $T=T_c$ (one expects the correct curve to display an upward cusp from knowledge of $I(T)$ at T_c); (c) $T=\frac{1}{2}T_c$.

(4.9) to $(1-\mu^2)\mathcal{E}/\mu$ instead of I, we again obtain a series of Stieltjes and the theorems of Ref. 1 apply.

In order to test this conjecture against available data, we note first of all that the condition

$$\operatorname{Re}[f(z)]\geq 0, \qquad |z|\leq 1 \tag{5.6}$$

is equivalent to the condition

$$|\sigma_0(z)|\leq 1, \qquad |z|\leq 1, \tag{5.7}$$

where

$$\sigma_0(z)=[f(z)-1]/[f(z)+1]. \tag{5.8}$$

Let us first consider a result of Nevanlinna.[14] The result is that it is necessary and sufficient for $|\sigma_0(z)|<1$ in $|z|<1$ for $|\sigma_n(0)|\leq 1$ for all n. We define

$$\sigma_n(z)=z^{-1}\left[\frac{\sigma_{n-1}(0)-\sigma_{n-1}(z)}{1-\sigma_{n-1}^*(0)\sigma_{n-1}(z)}\right], \tag{5.9}$$

[14] R. Nevanlinna, Ann. Acad. Sci. Fennicae, Ser. A13, 1 (1919).

where the asterisk denotes the complex conjugate. If $|\sigma_{n=1}(z)|<1$ for $|z|<1$, then the absolute value of the quantity within the brackets in (5.9) can easily be shown to be less than unity. By Schwartz's lemma, division by z does not increase the maximum modulus, and so $|\sigma_n(z)|<1$ also. Thus the inequalities are necessary. Nevanlinna also shows that they are sufficient. From (5.9) it follows that we may compute the first $\sigma_n(0)$ (we will abbreviate these by σ_n) directly from the power-series coefficients of $\sigma_0(z)$ through order z^n.

For the case $T=\infty$ ($\beta=0$), we have

$$(1-\mu^2)\mathcal{E}/\mu=\tfrac{1}{2}q(1-\mu)/(1+\mu), \tag{5.10}$$

where q is the lattice coordination number. One easily verifies property (5.6) or (5.7) in this case, thereby confirming the conjecture for that temperature. To test this conjecture further, we have tested it by means of the "σ test" described above for temperatures of $\frac{1}{2}T_c$, T_c, and $2T_c$ for the square, triangular, honeycomb, simple-cubic, body-centered-cubic, face-centered-cubic, and diamond lattices. In every case the test was met, and by a wide margin. The only σ's which were close to unity in absolute value were the σ_0's for $T=\frac{1}{2}T_c$. However, as

$$\sigma_0=(qu^{q-1}-1)/(qu^{q-1}+1), \tag{5.11}$$

where, if J is the exchange integral,

$$u=\exp(-4J/kT), \tag{5.12}$$

we necessarily have $1>\sigma_0>-1$ for ferromagnetic u (i.e., $0<u\leq 1$). All the other σ's are well below 1 in absolute value and are apparently *decreasing* as far as the series data of Sykes, Essam, and Gaunt[12] allow us to enquire. All the series evidence now available indicates that this conjecture is valid, and that the results of Padé-approximant bounds formed from the series of Stieltjes extracted from (5.5) are correct bounds to the configurational energy $\mathcal{E}$. Some sample results for the simple-cubic lattice are displayed in Fig. 6.

6. NONBOUNDING PROCEDURES FOR COMPUTING WITH THE μ SERIES TO THE ISING MODEL

Although the usual Padé-approximant procedure to the μ series does not give bounds in the region $H=0$, $T<T_c$, if we consider, for example, the spontaneous magnetization, they give consistent values to many decimal places very quickly. At $T=\frac{1}{2}T_c$, for the number of terms available,[12] no error is apparent in the eight figure results computed for loose-packed lattices and a fluctuation is apparent only in the sixth figure for close-packed lattices where fewer terms are known.

One word of caution, however, should be added at this point concerning the use of nonbounding summation procedures as applied to these series. If we consider the magnetization series for $T=T_c$, then we know

$I(1)=0$. However, straight Padé approximants give values (depending on lattice) of up to 0.75 with a consistency of a half a percent or so (square lattice, for example). As we know (theorem 3, Ref. 1) that the Padé approximant cannot converge to the wrong answer (regular point), we concluded that this is a reflection of very slow convergence. The Taylor series reflects similar slow convergence. Although the last term available (square lattice) μ^{13} is -4.1734×10^{-3}, the sum is still

$$I(1)\approx\sum_{j=0}^{13} m_j\approx 0.8148. \tag{6.1}$$

An additional 195 terms of the same size as the thirteenth would be required to reduce $I(1)$ to zero. The difficulty is that $H=0$, $T=T_c$ is known to be a singular point of I.

In the region $H=0$, $T>T_c$, Domb[15] has shown that, if Λ is the partition function per spin

$$\frac{\Lambda}{(1+\mu)}=\sum_{r=1}^{\infty}\frac{\psi_r(\mu)\zeta^r}{(1+\mu)^{2r}}, \tag{6.2}$$

where $\psi_r(\mu)$ is a polynomial of degree at most $2r$, and $\zeta=1-u$, a high-temperature variable. If we introduce the linear fractional transformation

$$\eta=\mu/(1+\mu), \tag{6.3}$$

then (6.2) becomes

$$\Lambda/(1+\mu)=\sum_{r=1}^{\infty}\phi_r(\eta)\zeta^r, \tag{6.4}$$

where the $\phi_r(\eta)$ are polynomials of degree $2r$ and have lower-order coefficients that vanish to progressively higher orders as r increases. Consequently, a Taylor expansion in η to order $2r$ will differ from the complete function by terms only of order $O(\zeta^{r+1})$. However, as the $[N,N]$ Padé approximants are invariant under (6.3) (theorem 1, Ref. 1), the same rapid convergence is to be expected in the high-temperature region from the $[N,N]$ summation of the μ series. Except for simple correction terms [as in (6.4)], the same type results are expected for other thermodynamic quantities.

As a practical matter then, we expect nonbounding summation procedures to produce good results throughout the H-T plane away from the critical point. They do however suffer from the drawback of deceptively slow convergence near the critical point, which can lead the unwary into serious underestimation of his error bounds.

A few remarks concerning analytic continuation in the complex μ plane are perhaps in order. Yang and Lee[9] have shown that the roots of the grand partition function lie on the unit circle in the μ plane. In the high-temperature limit they are all at $\mu=-1$. As the temperature decreases, they seem (Sec. 5) to move along the circle toward $\mu=+1$. They reach (in the limit of an infinite system) $\mu=+1$ at the critical temperature. For temperatures above T_c, $\mu=+1$ is a regular point and we can plainly continue to values of $\mu>1$. As the temperature approaches T_c from above, the poles move in on the point $\mu=+1$, completely closing the unit circle.

There is, however, no reason to suppose that just because the physical grand partition function has a solid wall of poles on the unit circle that it necessarily forms a natural barrier, any more than one would suppose

$$\tan^{-1}\left(\frac{x}{a}\right)=x\int_a^{\infty}\frac{dz}{x^2+z^2} \tag{6.5}$$

has a natural barrier for $-\infty<x^2<-a^2$, which is known to be untrue. That analytic continuation through the unit circle is sometimes possible can be illustrated in the case of the one-dimensional Ising model. According to Yang and Lee,[9] $g(\theta)$ in (4.10) for the free energy is given by

$$g(\theta)=(2\pi)^{-1}\frac{\sin\frac{1}{2}\theta}{(\sin^2\frac{1}{2}\theta-u^2)^{1/2}},\qquad \cos\theta<1-2u^2$$
$$=0,\qquad \cos\theta>1-2u^2. \tag{6.6}$$

However, we can rewrite (4.10) by using $\sin\frac{1}{2}\theta=[(\mu')^{1/2}-(\mu')^{-1/2}]/(2i)$ as

$$\frac{(F+mH)}{kT}=-(2\pi i)^{-1}\int_C \ln[1-\mu(\mu'+(\mu')^{-1})+\mu^2]\times\frac{(1-\mu')d\mu'/\mu'}{[1+2(2u^2-1)\mu'+\mu'^2]^{1/2}}, \tag{6.7}$$

where C is that part of the unit circle described in (6.6). We may now distort the contour in the μ' plane by Cauchy's theorem without changing the value of (6.7) and thus analytically continue the free energy onto an unphysical Riemann sheet. If the density function is smooth in higher dimensions, there is no reason to suppose that this sort of analytic continuation is not still possible. Only the endpoints of C are necessarily singular.

When $T=T_c$, the endpoints disappear. The most likely consequence is that a singular point remains at $\mu=+1$ [note the slow convergence of (6.1)] but that the $H=0^+$ and $H=0^-$ Riemann sheets tear apart. What is meant by Riemann sheets tearing apart is illustrated by (6.5) in the limit as $a\to0$. The result is $\pm\pi/2$ for all x where the sign is given by the sign of the real part of x. Clearly, analytic continuation of either of these branches is possible and does not give the other branch after the limit $a\to0$ has been taken. Experimentally (Padé approximation and ratios), it appears the singularities in the μ plane recede from the

[15] Equation (113), Sec. 5.4.3 of Ref. 10.

unit circle as the temperature is reduced below the critical temperature. These analytic continuations probably correspond to metastable states.

We conclude that every physical point in the μ-T plane is regular, except the critical point, and that standard Padé-approximation procedures to the μ series should be pleasantly convergent. However, the rate of convergence near the critical point is so deceptive that it is prudent to adopt a bounding summation procedure.

ACKNOWLEDGMENT

The author is happy to acknowledge several stimulating conversations with Dr. O. Penrose during the early phases of this work.

PHYSICAL REVIEW VOLUME 161, NUMBER 2 10 SEPTEMBER 1967

Electronic Theory of Phase Transitions in Ca, Sr, and Ba under Pressure*

A. O. E. Animalu

Division of Applied Physics, Stanford University, Stanford, California

(Received 2 March 1967; revised manuscript received 25 May 1967)

The transitions induced by temperature and pressure between the face-centered cubic (fcc) and the body-centered cubic (bcc) phases in the alkaline earth metals—Ca, Sr, and Ba—are analyzed by computing the differences in Gibbs free energy between the two phases in the framework of the nearly-free-electron and harmonic approximations. At the absolute zero of temperature and pressure, the observed fcc in Sr and bcc in Ba were found to have the lower internal energy; in Ca, however, identical analysis led to lower energy in bcc rather than the observed fcc. An fcc-bcc transition characterized by a change in the sign of the difference in free energies in Sr at 0°K was found at a critical pressure $P_c \sim 10$ kbar; at zero pressure, it was found at a critical temperature, $T_c \sim 150$°K. Both these results agree only qualitatively with the observed $P_c \sim 36$ kbar and $T_c \sim 830$°K. Ca and Ba, already bcc at absolute zero, showed no phase transition.

I. INTRODUCTION

RECENTLY, the changes in the electrical properties and the crystal structures of the alkaline-earth metals—Ca, Sr, and Ba—under pressure have attracted considerable attention. Efforts have, however, been focused primarily on the change of the resistance with pressure and its implications for the electronic band structure of the alkaline-earth metals.[1–5] From these studies, particularly the most recent extensive band-structure calculations in the face-centered cubic (fcc) and body-centered cubic (bcc) phases by Vasvari *et al.*,[3] it has become clear that the band structures near the Fermi surface are basically nearly-free-electron-like, as previously obtained by Harrison,[6] but do not have the simple form of *s*-*p* bands suggested by Mott[1] and Drickamer.[2] A computation of the high-pressure electrical resistance R by Vasvari and Heine[4] on the basis of the band structure in the fcc phase has shown that the high resistance and the negative $\partial R/\partial T$ in Ca and Sr can be understood in terms of a high-pressure semimetallic state with vanishingly small Fermi-surface area, rather than the semiconducting state obtained by Altmann and Cracknell.[5] In a slightly different vein from these developments, Jerome *et al.*[7] have recently developed a theory of the transition from semimetallic states characterized by small energy gaps or small band overlaps to a new state, which they have called excitonic insulator, and for which considerable studies have been reported in the Russian literature.

In this paper we shall make a quantitative study of the transitions in the crystal structures. In pure Ca and Sr, which crystallize in the fcc phase at 0°K, temperature-induced transitions to the bcc phase have been observed at 721° and 830°K, respectively.[8] Pressure-induced transitions were first reported by Bridgman,[9,10] and the phase boundaries measured by Jayaraman *et al.*[8] up to pressures of 45 kbar. In Ca, the transition temperature T_c rises with pressure: $\partial T/\partial P \sim +3.3$°C/bar at 1 atm; in Sr, T_c drops with pressure: $\partial T/\partial P \sim -10$°C/bar. In Ba, which crystallizes in the bcc phase at 0°K, temperature-induced transitions have not been observed, but a pressure-induced transition

* This work was supported by the Advanced Research Project Agency through the Center for Materials Research at Stanford University.

1 N. F. Mott, Phil. Mag. **13**, 989 (1966).
2 H. G. Drickamer, Solid State Phys. **17**, 1 (1965).
3 B. Vasvari, A. O. E. Animalu, and V. Heine, Phys. Rev. **154**, 535 (1967).
4 B. Vasvari and V. Heine (to be published).
5 S. L. Altmann and A. P. Cracknell, Proc. Phys. Soc. (London) **84**, 761 (1964).
6 W. A. Harrison, Phys. Rev. **131**, 2433 (1963).
7 D. Jerome, T. M. Rice, and W. Kohn (to be published).
8 A. Jayaraman, W. Klement, Jr., and G. C. Kennedy, Phys. Rev. **132**, 1620 (1963).
9 P. W. Bridgman, Proc. Am. Acad. Arts Sci. **72**, 187 (1938); **74**, 425 (1942); **81**, 169 (1952).
10 P. W. Bridgman, Phys. Rev. **60**, 351 (1941).